



[Home](#) › [Statistics](#) › [Microdata and TableBuilder](#) › [Available Microdata and TableBuilder](#)
› [Qualifications and work](#)

Microdata and TableBuilder: Qualifications and work

Contains TableBuilder and DataLab products

On this page

[Accessing the data](#)

[Data and file structure](#)

[Using DataLab](#)

[Using TableBuilder](#)

[Data downloads](#)

[Explanatory notes](#)

Accessing the data

The Qualifications and Work publication presents detailed information about the educational history of people and the relevance of each qualification to their working lives. Data collected for up to five qualifications includes the level and field, year of completion and whether the qualification was attained in Australia. Further information was collected on incomplete qualifications, cultural background and citizenship status. See [Qualifications and Work, Australia \(/statistics/people/education/qualifications-and-work/latest-release\)](#) for summary results, methodology and other information.

The following microdata products are available from this survey:

- TableBuilder - produce your own tables and graphs. TableBuilder is available for the following survey years: 2018/19, 2015 and 2010/11. See the TableBuilder data item lists included with this release.
- DataLab - expanded microdata is available in DataLab for the following survey years: 2018/19. See the microdata data item lists included with this release.

To apply for access to the DataLab, please follow the steps in [Apply for access \(/statistics/microdata-tablebuilder\)](#).

Data and file structure

Data available by level

The Qualifications and Work microdata is available across two levels.

1. Person Level
2. Qualification Level

The two levels are hierarchical. The first level relates to people and the second level relates to the qualifications of those people.

Visually, the person level file has one row of data for each respondent, containing all the Person level data items. The Qualification level contains up to five rows for each respondent, one for each of their qualifications.

The Person level contains characteristic data such as age, sex, marital status, employment, and educational status including the number of non-school qualifications completed and the level and field of highest educational attainment. This level also has information about the persons household, for example, the number of children present aged less than 15 and the overall household income. In addition, the Person level includes geographic data items such as state or territory of usual residence.

The Qualification Level contains details about each episode or completed qualification that each respondent has reported. The types of details available about each qualification include the level and field of the qualification, the year each qualification was started and completed, whether the qualification was completed in Australia and if employed, the relevance of each qualification to the person's current job.

A complete data item list can be accessed from the Data downloads section.

Unit identifiers

Every record on each level of the detailed microdata files are uniquely identified. For more information about the identifiers, see the Weights and Identifiers section of the relevant Data Item List.

Using DataLab

The DataLab environment allows real time access to detailed microdata from the Qualifications and Work Survey.

The DataLab is an interactive data analysis solution available for users to run advanced statistical analyses, for example, multiple regressions and structural equation modelling. The DataLab environment contains up-to-date versions of SPSS, Stata, SAS and R analytical languages. Controls in the DataLab have been put in place to protect the identification of individuals and organisations. All output from DataLab sessions is cleared by an ABS officer before it is released.

DataLab files

There are two Qualifications and Work Datalab files; a Person level file and a Qualification level file.

The Person level (O18QUAL) contains a range of data items relating to the respondent, including; demographic, employment, income, education, study and geographical details.

The Qualification level (O18QUALEP) contains specific details on each of the five highest qualifications that the respondent has obtained. The file is hierarchical with a unique record (row) for each of their qualifications.

For information about all of the data items available on these datasets please see the Microdata data item list in the Data downloads section.

DataLab test file

A test file is available in the Data downloads section for researchers to become familiar with the data structure and prepare code/programs before applying for or beginning a DataLab session.

For more information, including prerequisites for DataLab access, please see the [About the DataLab \(/datalab\)](#) page.

Counting units and weights

Weighting is the process of adjusting results from a sample survey to infer results for the total population. To do this, a 'weight' is allocated to each record. The weight is the value that indicates how many population units are represented by each sample unit.

If you are estimating the number of persons with certain characteristics (e.g. 'Number of non-school qualifications completed') the person level file and person level weight is used.

To estimate the total number of qualifications (e.g. the number of non-school qualifications completed) the qualification level and qualification weight is used. Qualifications are weighted according to the characteristics of the person who undertook the qualification, and therefore the weights for each qualification are the same as the weight for the person.

For example, if a person in the sample has a weight of 600 and that person has completed three non-school qualifications then the person represents 600 people in the total population and 1,800 qualifications.

The names of the weights and replicate weights can be found in the Weights and Identifiers tab of the data item list.

Combining person and qualification level data items

Every record on each file is uniquely identified. These Identifiers can be used to copy information from one level of the file to another. For example, demographic detail from the person level, such as age and sex, can be added to the qualification level (in a one to many merge) using the identifiers.

The names of the Identifiers can be found in the Weights and Identifiers tab of the data item list.

Qualification ordering flags

To assist with analysis, several variables have been created to help isolate and order qualification level data. These are detailed on the 'Five highest Qualifications' tab of the data item list.

Using a qualification order flag such as 'Highest to fifth highest non-school qualification' allows for one qualification to be selected from this level (such as the highest qualification). This will allow for estimates of persons to be produced, rather than an estimate of qualifications.

In summary, qualification level data items can be cross-tabulated with person level data items with or without qualification flags. Qualification flags should be included in analysis when a user wants information only about one particular qualification (e.g. the highest qualification or the most recent qualification), but should not be used in tables looking at all qualifications.

Using TableBuilder

For general information relating to the TableBuilder or instructions on how to use features of the TableBuilder product, please refer to the User Manual: TableBuilder (cat. no. 1406.0.55.005).

Information applicable to the Qualifications and Work TableBuilder, which enables users to understand, interpret and tabulate the data, is outlined below.

Confidentiality features in TableBuilder

In accordance with the Census and Statistics Act 1905, all data in TableBuilder are subjected to a confidentiality process before release. This confidentiality process is undertaken to avoid releasing information that may allow the identification of particular individuals, families, households, dwellings or businesses.

Processes used in TableBuilder to confidentialise records include the following:

- perturbation of data
- table suppression

Perturbation effects

To minimise the risk of identifying individuals in aggregate statistics, a technique is used to randomly adjust cell values. This technique is called perturbation. Perturbation involves small random adjustments of the statistics and is considered the most satisfactory technique for avoiding the release of identifiable statistics while maximising the range of information that can be released. These adjustments have a negligible impact on the underlying pattern of the statistics.

The introduction of these random adjustments result in tables not adding up. While some datasets apply a technique call additivity to give internally consistent results, additivity has not been implemented on this TableBuilder. As a result, randomly adjusted individual cells will be consistent across tables, but the totals in any table will not be the sum of the individual cell values. The size of the difference between summed cells and the relevant total will generally be very small.

Please be aware that the effects of perturbing the data may result in components being larger than their totals. This includes determining proportions.

Table suppression

Some tables generated within TableBuilder may contain a substantial proportion of very low counts within cells (excluding cells that have counts of zero). When this occurs, all values within the table are suppressed in order to preserve confidentiality. The following error message below is displayed at the bottom of the table when table suppression has occurred.

ERROR: The table has been suppressed as it is too sparse
ERROR: table cell values have been suppressed

Counting units and weights

Weighting is the process of adjusting results from a sample survey to infer results for the total population. To do this, a 'weight' is allocated to each record. The weight is the value that indicates how many population units are represented by each sample unit.

To produce estimates for the in-scope population you must use weight fields in your tables. In TableBuilder they can be found under the Summation Options category in the left hand pane under the applicable level. If you do not select a weight field, TableBuilder will apply 'Person weight' by default. This will give you estimates of the number of persons. To produce estimates of the number of qualifications you would have to add 'Qualification level weights' from the Qualification level to your table.

If you are estimating the number of persons with certain characteristics (e.g. 'Number of non-school qualifications completed') the weight listed under the category heading 'Person Level Weights' must be used. To estimate the number of qualifications (e.g. the number of non-school qualifications completed) the weight listed under 'Qualification level weights' must be used.

Qualification level data items are weighted according to the characteristics of the person who undertook the qualification, and therefore the weights for each qualification are the same as the weight for the person. For example, if a person in the sample has a weight of 600 and that person has completed three non-school qualifications then the person represents 600 people in the total population and 1,800 qualifications.

Selecting data items for cross-tabulation

The Person Level contains a range of data items detailing the characteristics of the respondent including some education variables. The Qualification Level contains data items about each of the qualifications that a respondent has obtained. The file is hierarchical with each respondent record potentially having multiple qualification records.

Cross-tabulating data items on the same level

Cross-tabulating data from the Person Level with other data items from the same level will produce data about people. For example, cross-tabulating the geographic variable 'State or territory of usual residence' by the 'Level of most recent non-school qualification' produces a table showing the number of people in each region by the most recent qualification they have obtained.

Cross-tabulating data from the Qualification Level with other data items from the same level will produce data about qualifications when using the Qualification Weight. For example, cross-tabulating 'Level of non-school qualification' by 'Whether completed qualification through an Australian institution' produces a table showing the number of qualifications

completed through an Australian institution. If a respondent has several qualifications, each of those qualifications is included in the table. If the same cross-tabulation is generated but using the Person weight instead of the Qualification weight, it produces a table showing the number of people who completed a non-school qualification through an Australian institution.

Using Qualification ordering data items and flags

To assist with analysis, several variables have been created to help isolate and order qualifications. The following shows the available Qualification level ordering data items and flag item.



By using a Qualification level ordering data item (e.g. 'Highest to fifth highest non-school qualification' and selecting the 'Highest non-school qualification' category) only one qualification for each respondent is included in a table. Selecting either the Person weight or the Qualification weight when using a Qualification ordering data item or flag will produce essentially the same result, any difference being the result of perturbation acting slightly differently when using the different weights.

Cross-tabulating Person Level by Qualification Level data items

Cross-tabulating data items from the Person Level with data items from the Qualification Level can produce data about people or qualifications depending on the weight being used. Caution should be used when Cross-tabulating a Qualification Level data item while using a Person weight as a person with multiple qualifications may have the same qualifications counted only once in a table.

Using a Qualification ordering data item or flag (as described in 'Using Qualification ordering data item and flags' above) may be worthwhile when cross-tabulating Person Level with Qualification Level data items as only one selected qualification will be included in the tabulation.

Cross-tabulating qualification level data items by person level data items using the person weight - When using a qualification ordering data item or flag

When using a Qualification ordering data item (e.g. 'Highest to fifth highest non-school qualification') and selecting the 'Highest non-school qualification' category' in a table that cross-tabulates a qualification level data item by a person level data item, either the Person or the Qualification weight can be used and the same output will be generated, with any difference being due to perturbation (see Perturbation Effects above). Restricting the table to a single qualification (e.g. highest non-school qualification) for each person in effect turns this into a person level data item, as TableBuilder only needs to read one row of data from the qualification level for each person.

Cross-tabulating qualification level data items by person level data items using the person weight - When NOT using a qualification ordering data item or flag

When a Qualification ordering data item is not used, TableBuilder will read each row of data from the qualification level for each person. In this case, TableBuilder effectively calculates the tabulation as a 'multi-response' table (i.e. the same person can be counted more than once), but it counts the same categories of information about different qualifications only once. It treats them as 'one or more occurrences' of that category. For example if a respondent completed three qualifications, and the respondent is currently working in the field as one of these qualifications but not the other two then the person would be counted in each column of the data item 'Currently working in the same field as main field of non-school qualification'.

Therefore, in these particular types of tabulations, components of the table will not add to the total number of persons (as persons can be counted more than once), but the total will be the correct count of persons as TableBuilder calculates the total in such a way that each person is only counted once. An example table below (without a qualification raking data item) for 'Currently working in the same field as main field of non-school qualification' shows results for all qualifications for a person, so they can appear in both columns:

Persons ⓘ	Persons #		
Age of person (10 year age groups) ⓘ ∑ ☰	age 15-64		
Currently working in the same field as main field of non-school qualification - Qual level ⓘ ∑ ☰	Currently working in the same field as main field of study of qualification	Not currently working in the same field as main field of study of qualification	Total
≤ Labour force status ⓘ ∑ ☰	Employed	Employed	Employed
Level of non-school qualification (ASCED) - Qualification level ⓘ ⓘ ⓘ ☰	☰	☰	☰
<u>Postgraduate Degree</u>	753.1	274.1	995.2
<u>Graduate Diploma and Graduate Certificate</u>	524.5	190.5	683.9
<u>Bachelor Degree</u>	2,507.8	1,330.9	3,701.9
<u>Advanced Diploma and Diploma</u>	1,111.2	761.5	1,805.6
<u>Certificate III and IV</u>	2,139.6	1,276.7	3,213.8
<u>Certificate I and II</u>	212.2	245.3	432.2
<u>Certificate n.f.d.</u>	79.4	42.4	119.9
<u>Level not determined</u>	340.5	197.2	522.6
Total	6,172.5	3,647.9	8,443.2
▼ Annotation Descriptions			
# (000's)			

The following table shows the same data item for the 'Highest qualification' only by using the ordering data item 'Highest to fifth highest non-school qualification', so people only appear once, only in either column and consequently columns add to totals (taking perturbation into account, see Perturbation Effects above).

Persons ⓘ	Persons #		
Age of person (10 year age groups) ⓘ ∑ ☰	age 15-64		
Currently working in the same field as main field of non-school qualification - Qual level ⓘ ⓘ ⓘ ☰	Currently working in the same field as main field of study of qualification	Not currently working in the same field as main field of study of qualification	Total
≤ Labour force status ⓘ ∑ ☰	Employed	Employed	Employed
Highest to fifth highest non-school qualification ⓘ ⓘ ⓘ ☰	Highest qualification	Highest qualification	Highest qualification
Level of non-school qualification (ASCED) - Qualification level ⓘ ⓘ ⓘ ☰	☰	☰	☰
<u>Postgraduate Degree</u>	740.3	255.7	995.2
<u>Graduate Diploma and Graduate Certificate</u>	418.0	127.8	545.0
<u>Bachelor Degree</u>	1,716.9	840.9	2,554.6
<u>Advanced Diploma and Diploma</u>	794.2	555.3	1,345.7
<u>Certificate III and IV</u>	1,643.2	832.5	2,473.7
<u>Certificate I and II</u>	112.3	108.4	223.7
<u>Certificate n.f.d.</u>	36.7	* 13.9	52.2
<u>Level not determined</u>	165.1	89.0	255.0
Total	5,619.0	2,823.5	8,443.2
▼ Annotation Descriptions			
# (000's)			

In summary, qualification level data items can be cross-tabulated with person level data items with or without Qualification ordering data items or flags. Qualification ordering data items and their relevant category should be included in tables when a user wants information only about one particular qualification (e.g. the highest qualification or the most

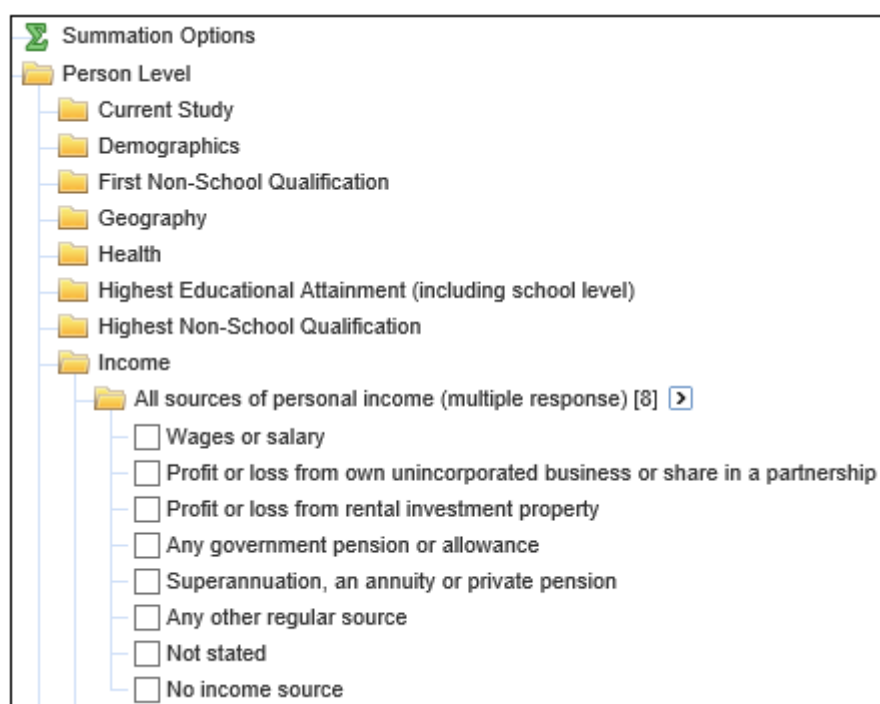
recent qualification), but should not be used in tables looking at all qualifications.

Zero value cells

Tables generated from sample surveys will sometimes contain cells with zero values because no respondents that satisfied the parameters of the cell were in the survey. This is despite there being people in the population with those characteristics. That is, the cell may have had a value above zero if all persons in scope of the survey had been enumerated. This is an example of sampling variability which occurs with all sample surveys. Relative Standard Errors cannot be generated for zero cells.

Multi-response data items

A number of the survey's data items allow respondents to report more than one response. These are referred to as 'multi-response data items'. When a multi-response data item is tabulated, a person is counted against each response they have provided (e.g. a person who responds 'employee income' and 'unincorporated income' and 'government pensions and allowances' will be counted once in each of these three categories).



As a result, each person in the appropriate population is counted at least once, and some persons are counted multiple times. Therefore, the total for a multi-response data item will be less than or equal to the sum of its components.

Not applicable categories

Most data items include a 'Not applicable' category. The 'Not applicable' category comprises those respondents who were not asked a particular question(s) and hence are not

applicable to the population to which the data item refers. 'Not applicable' for a lot of our derived items are those people who are not in the population, such as people without qualifications or who are not employed. For example, the Not applicable category for the data item 'Level of highest non-school qualification (ASCED)' comprises those people who have not completed a non-school qualification and for 'Full-time or part-time status of employment in all jobs or businesses' it comprises those people who are not employed. The population and the classification value of the 'Not applicable' category (where relevant) is shown in the data item list (see the Data Item List in the Data downloads section).

Table populations

The population relevant to each data item is shown in the data item list and should be considered when extracting and analysing the microdata. The actual population count for each data item is equal to the total cumulative frequency minus the 'Not applicable' category.

Data downloads

Data cubes

⬇ Download all (556.83 KB)

Q&W 2018-19 Microdata data item list

⬇ [Download XLS](#)
[1.8 MB]

Q&W 2015 Microdata data item list

⬇ [Download XLS](#)
[565 KB]

L&W 2010-11 Microdata data item list

⬇ [Download XLS](#)
[275.5 KB]

Detailed Microdata test file

⬇ [Download ZIP](#)
[9.19 MB]

Explanatory notes

Show all

See [Qualifications and Work, Methodology \(/statistics/people/education/qualifications-and-](#)

[work/latest-release#methodology\)](#) for information on:

- How the data is collected
- How the data is processed
- Key education concepts
- Comparing the data
- How the data is released
- Accuracy
- Glossary
- Abbreviations

Previous catalogue number

This release previously used catalogue number 4235.0.55.001.